5)4/)
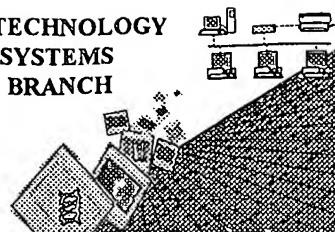
RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information
Center (STIC) detected errors when processing the following computer readable
form:

Application Serial Number: _10/048,146_
Source: _Public_
Date Processed by STIC: _2/13/2002_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE
   APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A
   NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
      http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is
a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the
United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm> , EFS Submission
   User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
   Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
             Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
   2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
   Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

PCT10

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/048,146

DATE: 02/13/2002
TIME: 10:07:52

see p.5

Input Set : A:\EP.txt
Output Set: N:\CRF3\02132002\J048146.raw

```
 3 <110> APPLICANT: The Government of the United States of America as represented by the
 4       Secretary of the Department of Health and Human Services, Centers for Disease
 5       Control and Prevention
 6       Tsang, Victor
 7       Greene, Ryan
 8       Wilkins, Patricia
 9       Hancock, Kathy
11 <120> TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING LARVAL TAENIA SOLIUM
13 <130> FILE REFERENCE: 6395-62068
15 <140> CURRENT APPLICATION NUMBER: US/10/048,146
15 <141> CURRENT FILING DATE: 2002-01-25
15 <150> PRIOR APPLICATION NUMBER: US 60/147,318
16 <151> PRIOR FILING DATE: 1999-08-03
18 <150> PRIOR APPLICATION NUMBER: PCT/US00/21173
19 <151> PRIOR FILING DATE: 2000-08-03
21 <160> NUMBER OF SEQ ID NOS: 9
23 <170> SOFTWARE: PatentIn version 3.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 2153
27 <212> TYPE: DNA
28 <213> ORGANISM: Taenia solium
30 <220> FEATURE:
31 <221> NAME/KEY: CDS
32 <222> LOCATION: (145)..(531)
33 <223> OTHER INFORMATION:
36 <400> SEQUENCE: 1
37 ctgcagtgaa gttgacaagt agttgaccat ttacggaaca tcaatggagg acactttggt      60
39 agggaaagca tacgataaac ataaaccaat gctggttata taagagacga tctcggctac     120
41 acttgtaact gaacaacctg taga atg cgt gcc tac att gtg ctt ctc gct        171
42                               Met Arg Ala Tyr Ile Val Leu Leu Ala
43                                1               5
45 ctc act gtt ttc gta gtg acg gtg tcg gcc gag tgg gtg ccc att tcg       219
46 Leu Thr Val Phe Val Val Thr Val Ser Ala Glu Trp Val Pro Ile Ser
47 10              15              20              25
49 agt gtc cac ata gcc tca tgc aaa agc cac tac atg ttc caa tta aaa       267
50 Ser Val His Ile Ala Ser Cys Lys Ser His Tyr Met Phe Gln Leu Lys
51              30              35              40
53 cgc ttt ttt gcc ttt agg aaa aac aaa ccg aaa gat gtt gca aat agt       315
54 Arg Phe Phe Ala Phe Arg Lys Asn Lys Pro Lys Asp Val Ala Asn Ser
55              45              50              55
57 acg aaa aaa ggg ata gaa tat gtc cac gaa ttc ttc cac gaa gac ccg       363
58 Thr Lys Lys Gly Ile Glu Tyr Val His Glu Phe Phe His Glu Asp Pro
59              60              65              70
```

```
61 att ggt aaa caa att gct caa ctc gca aag gaa tgg aag gaa gca atg          411
62 Ile Gly Lys Gln Ile Ala Gln Leu Ala Lys Glu Trp Lys Glu Ala Met
63      75                  80                  85
65 ttg gaa ggt agg ttt tgg tgt ttt ctg tca gaa gaa aat tat cta ttc          459
66 Leu Glu Gly Arg Phe Trp Cys Phe Leu Ser Glu Glu Asn Tyr Leu Phe
67 90                  95                  100                 105
69 att cat cta gac aaa ggc aaa ata cgg acg tca ctg gtt gag cac tgc          507
70 Ile His Leu Asp Lys Gly Lys Ile Arg Thr Ser Leu Val Glu His Cys
71                 110                 115                 120
73 aaa ggt cct aag aaa aaa act gct taacttgtca actttcatgc gttcttctct          561
74 Lys Gly Pro Lys Lys Lys Thr Ala
75                 125
77 tcactaataa atgctcatta ataagaaagc tgccttttgc aagatcaacg agggccatag       621
79 actgtgaggg ttatagccta aggttatggg gtgaaatgag ataggaattg agcatttgag       681
81 aagttactaa tttaaattga aagccgcatt tcttctgcaa ttgacgtgtg atggttagcg       741
83 aaaccaagtg aagcacgacc tcttgagtcg tttcaacagc cgccagtggt ttcaccagtg       801
85 gcttcaccag tgggtagact ggtttgtcac acatgcgagg tacggtcaga gggctaacag       861
87 gtgtggtgga ggggccaaca cgtgtaagac aagcagttcc ccttctctgt cgtgaggcac       921
89 actcagcacc cacctcgttt acttctccct tgacgactgt aatgcatttg gggtcaccat       981
91 gcccccgcca agttgaaggc actgatgaca tttgtaccat atcaccgata agtattaact      1041
93 cttccacttc ccagattttg aggtcaggcg atcctactga ctcggtgtag ccccatggtg      1101
95 gtccatgctc tgcaccattc gctgttcagt ggagcatcca cctagacggc caaccaatct      1161
97 cgcctccctt ctcctgtgct caagatgtgc gtcggtgaga tttggagggt ctgatcacca      1221
99 tactaaccac gtaggtttca tcatctctaa gaagcaccac ttcttgaggt cgcattgtgt      1281
101 accaccagcc ggtgtaatca agagtgactt tcgcgtcacc cctaagaagg ctatagatct     1341
103 gcaagtcagc gcaatagctt cagccatgct gactaaaatg tgtaagggac cagtagctct     1401
105 agcccaacac aagtggagct aataatgggc ttccccagat acatgaatcc caaatcggtg     1461
107 agcatgggcc atgaatatgg ccttctgagt cttccttgaa tgcaaacgaa ggcatagcac     1521
109 gagggtagga tgagtgtaca gaaaacagcg aggcaacgaa tctactggca tggccctgat     1581
111 gccaccccgc ccagctaggg tagtttggcc acctcagtcc ttaatcgaat gcggcagtca     1641
113 gaacaaacaa agtattacat agccacactc ttcttttgag cgtcgtcctc gacgctcctt     1701
115 tcgacacacc tcccgcatca gccaccacaa agtaatcagt actggggaga cacccacgag     1761
117 ctaaccgtgc cagtcatgga aaatttgacg gcaactgagg agatgcctga ccccctttgg     1821
119 cagttcgaat gctgcccgtg gtcaaactcc tgcatcagcc atcacctacg attcaaacat     1881
121 cctagtcgcc aaattttcgt gaaccctcta aaattttcgt gcactctcaa gacacttcca     1941
123 actgacttag agctttttca tttggtgaga acacgtaaaa gcttcaagta aacaacaggc     2001
125 aacgatttca ctttgatgct ctcaccatca attctcttgt atgtgccacc accttaaacc     2061
127 ctccctgacc acttccactc tctctctctc cctaaataac aacacttgga agcatgaatg     2121
129 gtgtctgtca agttacacc cctagactgc ag                                     2153
132 <210> SEQ ID NO: 2
133 <211> LENGTH: 129
134 <212> TYPE: PRT
135 <213> ORGANISM: Taenia solium
137 <400> SEQUENCE: 2
139 Met Arg Ala Tyr Ile Val Leu Leu Ala Leu Thr Val Phe Val Val Thr
140 1               5                   10                  15
143 Val Ser Ala Glu Trp Val Pro Ile Ser Ser Val His Ile Ala Ser Cys
144             20                  25                  30
147 Lys Ser His Tyr Met Phe Gln Leu Lys Arg Phe Phe Ala Phe Arg Lys
```

RAW SEQUENCE LISTING   DATE: 02/13/2002
PATENT APPLICATION: US/10/048,146   TIME: 10:07:52

            Input Set : A:\EP.txt
            Output Set: N:\CRF3\02132002\J048146.raw

```
148        35                    40                    45
151 Asn Lys Pro Lys Asp Val Ala Asn Ser Thr Lys Lys Gly Ile Glu Tyr
152     50                    55                    60
155 Val His Glu Phe Phe His Glu Asp Pro Ile Gly Lys Gln Ile Ala Gln
156 65                    70                    75                    80
159 Leu Ala Lys Glu Trp Lys Glu Ala Met Leu Glu Gly Arg Phe Trp Cys
160                   85                    90                    95
163 Phe Leu Ser Glu Glu Asn Tyr Leu Phe Ile His Leu Asp Lys Gly Lys
164             100                   105                   110
167 Ile Arg Thr Ser Leu Val Glu His Cys Lys Gly Pro Lys Lys Lys Thr
168         115                   120                   125
171 Ala
175 <210> SEQ ID NO: 3
176 <211> LENGTH: 298
177 <212> TYPE: DNA
178 <213> ORGANISM: Taenia solium
180 <220> FEATURE:
181 <221> NAME/KEY: CDS
182 <222> LOCATION: (3)..(224)
183 <223> OTHER INFORMATION:
186 <400> SEQUENCE: 3
187 ta ttc gta gtg gcg gtt tcg gcc gag aaa aac aaa ccg aag tgt gat       47
188    Phe Val Val Ala Val Ser Ala Glu Lys Asn Lys Pro Lys Cys Asp
189    1                   5                   10                  15
191 gca aat agt act aag aaa gag ata gaa tat atc cac aat tgg ttt ttc      95
192 Ala Asn Ser Thr Lys Lys Glu Ile Glu Tyr Ile His Asn Trp Phe Phe
193                 20                    25                    30
195 cat gat gac ccg att gga aaa caa att gct caa ctc gca aag gac tgg     143
196 His Asp Asp Pro Ile Gly Lys Gln Ile Ala Gln Leu Ala Lys Asp Trp
197             35                    40                    45
199 aat gaa aca gtg cag gaa gcc aaa ggc aaa ttt tgg gcg tca ctg gct     191
200 Asn Glu Thr Val Gln Glu Ala Lys Gly Lys Phe Trp Ala Ser Leu Ala
201         50                    55                    60
203 gag tac tgc aga ggt ctg aag aac aaa act gct taacttgtca actttcatgc   244
204 Glu Tyr Cys Arg Gly Leu Lys Asn Lys Thr Ala
205     65                    70
207 gttcttctct tcaccaataa atgctgatta acaagaaaaa aaaaaaaaaa aaaa         298
210 <210> SEQ ID NO: 4
211 <211> LENGTH: 74
212 <212> TYPE: PRT
213 <213> ORGANISM: Taenia solium
215 <400> SEQUENCE: 4
217 Phe Val Val Ala Val Ser Ala Glu Lys Asn Lys Pro Lys Cys Asp Ala
218 1                   5                   10                  15
221 Asn Ser Thr Lys Lys Glu Ile Glu Tyr Ile His Asn Trp Phe Phe His
222             20                    25                    30
225 Asp Asp Pro Ile Gly Lys Gln Ile Ala Gln Leu Ala Lys Asp Trp Asn
226         35                    40                    45
229 Glu Thr Val Gln Glu Ala Lys Gly Lys Phe Trp Ala Ser Leu Ala Glu
```

```
230     50                 55                      60
233 Tyr Cys Arg Gly Leu Lys Asn Lys Thr Ala
234 65                 70
237 <210> SEQ ID NO: 5
238 <211> LENGTH: 294
239 <212> TYPE: DNA
240 <213> ORGANISM: Taenia solium
242 <220> FEATURE:
243 <221> NAME/KEY: CDS
244 <222> LOCATION: (3)..(221)
245 <223> OTHER INFORMATION:
248 <400> SEQUENCE: 5
249 tt ttc gta gtg gcg gtg tcg gcc gag gaa act aaa cca gag gac gtg        47
250    Phe Val Val Ala Val Ser Ala Glu Glu Thr Lys Pro Glu Asp Val
251    1              5              10              15
253 gta aag aat att aag aaa ggg atg gaa gtt gtc tac aaa ttt ttc tac       95
254 Val Lys Asn Ile Lys Lys Gly Met Glu Val Val Tyr Lys Phe Phe Tyr
255              20              25              30
257 gaa gac ccg ttg gga aag aaa ata gct caa ctc gca aag gac tgg aag      143
258 Glu Asp Pro Leu Gly Lys Lys Ile Ala Gln Leu Ala Lys Asp Trp Lys
259          35              40              45
261 gaa gca atg ttg gaa gcc aga agc aaa gtg cgg gcg tca ctg gct gag      191
262 Glu Ala Met Leu Glu Ala Arg Ser Lys Val Arg Ala Ser Leu Ala Glu
263        50              55              60
265 tac atc aga ggt ctc aag aac gaa gct gct taacttgtca actttcatgc        241
266 Tyr Ile Arg Gly Leu Lys Asn Glu Ala Ala
267        65              70
269 gttcttctct tcactaataa atgctcatta ataagaaaaa aaaaaaaaaa aaa           294
272 <210> SEQ ID NO: 6
273 <211> LENGTH: 73
274 <212> TYPE: PRT
275 <213> ORGANISM: Taenia solium
277 <400> SEQUENCE: 6
279 Phe Val Val Ala Val Ser Ala Glu Glu Thr Lys Pro Glu Asp Val Val
280 1              5              10              15
283 Lys Asn Ile Lys Lys Gly Met Glu Val Val Tyr Lys Phe Phe Tyr Glu
284              20              25              30
287 Asp Pro Leu Gly Lys Lys Ile Ala Gln Leu Ala Lys Asp Trp Lys Glu
288          35              40              45
291 Ala Met Leu Glu Ala Arg Ser Lys Val Arg Ala Ser Leu Ala Glu Tyr
292        50              55              60
295 Ile Arg Gly Leu Lys Asn Glu Ala Ala
296 65                 70
299 <210> SEQ ID NO: 7
300 <211> LENGTH: 6
301 <212> TYPE: PRT
302 <213> ORGANISM: Taenia solium
304 <400> SEQUENCE: 7
306 Ile Ala Gln Leu Ala Lys
```

```
307  1                    5
310  <210> SEQ ID NO: 8
311  <211> LENGTH: 24
312  <212> TYPE: PRT
313  <213> ORGANISM: Taenia solium
315  <220> FEATURE:
316  <221> NAME/KEY: variant
317  <222> LOCATION: (7)..(8)
318  <223> OTHER INFORMATION: Amino acid at position 7 may also be valine
321  <220> FEATURE:
322  <221> NAME/KEY: site
323  <222> LOCATION: (21)..(22)
324  <223> OTHER INFORMATION: Asparagine at position 21 is an amino acid insertion
327  <220> FEATURE:
328  <221> NAME/KEY: variant
329  <222> LOCATION: (14)..(15)
330  <223> OTHER INFORMATION: Amino acid at position 14 may also be glycine
333  <220> FEATURE:
334  <221> NAME/KEY: variant
335  <222> LOCATION: (18)..(19)
336  <223> OTHER INFORMATION: Amino acid at position 18 may also be valine
339  <220> FEATURE:
340  <221> NAME/KEY: variant
341  <222> LOCATION: (19)..(20)
342  <223> OTHER INFORMATION: Amino acid at position 19 may also be histidine
345  <220> FEATURE:
346  <221> NAME/KEY: variant
347  <222> LOCATION: (20)..(21)
348  <223> OTHER INFORMATION: Amino acid at position 20 may also be arginine
351  <400> SEQUENCE: 8
353  Lys Asn Lys Pro Lys Asp Asp Ala Ala Ser Thr Lys Lys Glu Ile Glu
354  1                 5                  10              15
357  Tyr Ile Trp His Asn Phe Phe Phe
358  20
361  <210> SEQ ID NO: 9
362  <211> LENGTH: 13
363  <212> TYPE: PRT
364  <213> ORGANISM: Taenia solium
366  <220> FEATURE:
367  <221> NAME/KEY: variant
368  <222> LOCATION: (5)..(6)
369  <223> OTHER INFORMATION: Amino acid at position 5 may also be isoleucine
372  <220> FEATURE:
373  <221> NAME/KEY: variant
374  <222> LOCATION: (12)..(13)
375  <223> OTHER INFORMATION: Amino acid at position 12 may also be aspartic acid
378  <220> FEATURE:
379  <221> NAME/KEY: variant
380  <222> LOCATION: (7)..(9)
```

[handwritten annotations:] "Asp" can only represent itself. Use Xaa, instead, and explain in (2207-(223) section

use Xaa

Xaa

Xaa

Xaa

Xaa

Xaa

Please correct any subsequent variants

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/048,146

DATE: 02/13/2002
TIME: 10:07:53

Input Set : A:\EP.txt
Output Set: N:\CRF3\02132002\J048146.raw

L:15 M:270 C: Current Application Number differs, Replaced Current Application No
L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date